

**Problems
in
Quantitative Linguistics
5**

by

Gabriel Altmann

**2015
RAM-Verlag**

Studies in quantitative linguistics

Editors

| | |
|-----------------|--|
| Fengxiang Fan | (fanfengxiang@yahoo.com) |
| Emmerich Kelih | (emmerich.kelih@univie.ac.at) |
| Reinhard Köhler | (koehler@uni-trier.de) |
| Ján Mačutek | (jmacutek@yahoo.com) |
| Eric S. Wheeler | (wheeler@ericwheeler.ca) |

1. U. Strauss, F. Fan, G. Altmann, *Problems in quantitative linguistics 1*. 2008, VIII + 134 pp.
2. V. Altmann, G. Altmann, *Anleitung zu quantitativen Textanalysen. Methoden und Anwendungen*. 2008, IV+193 pp.
3. I.-I. Popescu, J. Mačutek, G. Altmann, *Aspects of word frequencies*. 2009, IV +198 pp.
4. R. Köhler, G. Altmann, *Problems in quantitative linguistics 2*. 2009, VII + 142 pp.
5. R. Köhler (ed.), *Issues in Quantitative Linguistics*. 2009, VI + 205 pp.
6. A. Tuzzi, I.-I. Popescu, G. Altmann, *Quantitative aspects of Italian texts*. 2010, IV+161 pp.
7. F. Fan, Y. Deng, *Quantitative linguistic computing with Perl*. 2010, VIII + 205 pp.
8. I.-I. Popescu et al., *Vectors and codes of text*. 2010, III + 162 pp.
9. F. Fan, *Data processing and management for quantitative linguistics with Foxpro*. 2010, V + 233 pp.
10. I.-I. Popescu, R. Čech, G. Altmann, *The lambda-structure of texts*. 2011, II + 181 pp.
11. E. Kelih et al. (eds.), *Issues in Quantitative Linguistics Vol. 2*. 2011, IV + 188 pp.
12. R. Čech, G. Altmann, *Problems in quantitative linguistics 3*. 2011, VI + 168 pp.
13. R. Köhler, G. Altmann (eds.), *Issues in Quantitative Linguistics Vol 3*. 2013, IV + 403 pp.
14. R. Köhler, G. Altmann, *Problems in Quantitative Linguistics Vol. 4*. 2014, VIII+148 pp.
15. Best, K.-H., Kelih, E. (eds.), *Entlehnungen und Fremdwörter: Quantitative Aspekte*. 2014. VI + 163 pp.
16. I.-I. Popescu, K.-H. Best, G. Altmann, G. *Unified Modeling of Length in Language*. 2014, VIII + 123 pp.
17. G. Altmann, R. Čech, J. Mčutek, L. Uhlřřov (eds.), *Empirical Approaches to Text and Language Analysis dedicated to Luděk Hřřebek on the occasion of his 80th birthday*. 2014. VI + 231 pp.

18. M. Kubát, V. Matlach., R. Čech., *QUITA Quantitative Index Text Analyzer*. 2014, VII + 106
19. K.-H. Best, *Studien zur Geschichte der Quantitativen Linguistik*. 2015. III + 158 pp.
20. G. Altmann, *Problems in Quantitative Linguistics Vol. 5*. 2015, III+146 pp.

ISBN: 978-3-942303-33-0

© Copyright 2015 by RAM-Verlag, D-58515 Lüdenscheid

RAM-Verlag
Stüttinghauser Ringstr. 44
D-58515 Lüdenscheid
Germany
RAM-Verlag@t-online.de
<http://ram-verlag.de>

Preface

The present volume is a continuation of the series dedicated to all linguists who want to solve linguistic problems in a non-classical way. Elementary knowledge of statistics is a necessary condition, however, even a collection of data in the prescribed way could be helpful for solving some problems. The comparisons, tests, finding a function or distribution can be made by a statistician but the linguistic background knowledge must be furnished by the linguist.

The volume is appropriate especially for those who try to enter the field of quantitative linguistics and seek the door leading to elementary problems.

The present volume contains 90 problems. To each problem some references are recommended but the reader can solve them in his own way. Unfortunately, qualitative linguistics contains many concepts and classifications rooted in opinions and leading to different descriptions. In the present volume the reader is forced to perform tests which corroborate or reject the primary concept formation and force him to create new data based on different definitions, concepts, criteria etc. The basic requirement is the testing of everything one says.

It is recommended to publish the results in a quantitative linguistics journal. In any case, all numbers should be presented in order to give other linguists the possibility of testing other hypotheses or to subsume the accepted results in a deeper theory.

Gabriel Altmann

Contents

| | |
|--|-----------|
| 1. General Problems | 1 |
| 1.1. The problem of the problem | 1 |
| 1.2. Hierarchies in language | 2 |
| 1.3. Diversification of a language family | 4 |
| 1.4. Formal diversification | 6 |
| 1.5. Ways into the depth | 8 |
| 1.6. Length levels | 10 |
| 1.7. Hypotheses | 12 |
| 1.8. Distance and similarity | 13 |
| 1.9. Irregularity | 15 |
| 1.10. Problem continuation | 16 |
| | |
| 2. Text | 18 |
| 2.1. Frequency motifs | 18 |
| 2.2. Runs | 19 |
| 2.3. Sequences in text | 21 |
| 2.4. Frequency sequences | 25 |
| 2.5. The world view of language | 26 |
| 2.6. Climax types | 27 |
| 2.7. Continuous modeling of sentence length | 29 |
| 2.8. Thematic concentration 1 | 30 |
| 2.9. Thematic concentration 2 | 31 |
| 2.10. Thematic concentration 3 | 32 |
| 2.11. Thematic concentration 4 | 33 |
| 2.12. Denotative-connotative concentration | 34 |
| 2.13. Text compactness | 35 |
| 2.14. Conceptual inertia of texts 1 | 36 |
| 2.15. Conceptual inertia of texts 2 | 37 |
| 2.16. Conceptual inertia of texts 3 | 38 |
| 2.17. Adjective-verb ratio and text indicators | 39 |
| 2.18. Unified word length modeling | 41 |
| 2.19. Poetic and rhetoric figures 1 | 42 |
| 2.20. Poetic and rhetoric figures 2 | 44 |
| 2.21. The world view of a writer | 45 |
| 2.22. Adjectives in texts | 47 |
| 2.23. Adjectival motifs | 49 |
| 2.24. Stylistic centrality | 50 |
| 2.25. Text sorts | 51 |

| | | |
|-----------|--|-----------|
| 2.26. | Nominativity vs. predicativity 1 | 52 |
| 2.27. | Nominativity vs. predicativity 2 | 53 |
| 2.28. | Predication | 54 |
| 2.29. | Sentence length | 55 |
| 2.30. | Sequence of valencies | 57 |
| 3. | Grammar | 60 |
| 3.1. | Adnominal modifiers: Symmetry | 60 |
| 3.2. | Morphological complexity 1 | 61 |
| 3.3. | Morphological complexity 2 | 62 |
| 3.4. | Morphological changes and frequency | 63 |
| 3.5. | Word classes | 64 |
| 3.6. | Parts-of-speech distribution | 66 |
| 3.7. | Parts-of-speech homogeneity | 68 |
| 3.8. | Clause centrality | 69 |
| 3.9. | Clause types | 71 |
| 3.10. | Clause length | 72 |
| 3.11. | Topic – comment | 74 |
| 3.12. | Study of adverbials 1 | 77 |
| 3.13. | Study of adverbials 2 | 80 |
| 3.14. | Study of adverbials 3 | 81 |
| 3.15. | Grammatical categories | 82 |
| 3.16. | Reduplication | 83 |
| 3.17. | Sequences of syntactic constituents | 85 |
| 3.18. | Noun phrase | 87 |
| 3.19. | Syntactic tags as NP components | 88 |
| 3.20. | Word class specification | 89 |
| 4. | Stage play | 92 |
| 4.1. | Stage play: 1. Sentence length | 92 |
| 4.2. | Stage play: 2. Speech acts | 93 |
| 4.3. | Stage play: 3. Sequences of illocutive speech acts | 95 |
| 4.4. | Stage play: 4. Transition matrix of speech acts | 97 |
| 4.5. | Stage play: 5. Polysemy in the speech of persons | 98 |
| 4.6. | Stage play: 6. Distant reaction | 99 |
| 4.7. | Stage play: 7. Aggregation of speech acts | 100 |
| 4.8. | Stage play: 8. Act compactness | 101 |
| 4.9. | Stage play: 9. Verb activity | 102 |

| | |
|---|------------|
| 5. Phonemics and script | 104 |
| 5.1. Phonological complexity | 104 |
| 5.2. Script motifs | 106 |
| 5.3. Phonic similarity of words in proverbs | 107 |
| 5.4. A variant of the Skinner hypothesis | 109 |
| 5.5. Syllable complexity | 110 |
| 5.6. Euphony | 111 |
| 5.7. Distribution of syllable types | 112 |
| 5.8. Phonetic symbolism | 114 |
| | |
| 6. Semantics | 117 |
| 6.1. Abstractness of nouns | 117 |
| 6.2. Abstractness in the text | 118 |
| 6.3. Polysemy | 119 |
| 6.4. Measurement of verb activity | 121 |
| 6.5. Word class specification | 123 |
| 6.6. Metaphor | 124 |
| | |
| 7. Other problems | 127 |
| 7.1. Morphological motifs | 127 |
| 7.2. Sentence motifs | 128 |
| 7.3. Borrowings | 129 |
| 7.4. Syllabic word length | 130 |
| 7.5. Clause length | 132 |
| 7.6. Word length and number of compounds | 133 |
| 7.7. Word frequency and number of compounds | 134 |
| | |
| Author register | 136 |
| Subject register | 143 |